41 PRTS

Field of the invention

The present invention is based on the finding that two fimbrial operons, the *saf* operon and the *tcf* operon, are specific for *Salmonella enterica* subspecies 1 bacteria and therefore have therapeutic use. Due to their specificity they can be used to provide vaccines against *Salmonella enterica* subspecies I as well as for detection of *Salmonella enterica* subspecies I. The *saf* operon is specific for all *Salmonella enterica* subspecies 1 bacteria and the *tcf* operon is specific for the serovar Typhi of *Salmonella enterica* subspecies 1.

All or part of the DNA-sequences of the genes encoding these proteins can be used as active agents in a vaccine against diseases caused by the *Salmonella enterica* subspecies I bacterial strains or for detection of said bacterial strains.

The present invention also relates to methods of isolating these fimbrial proteins, to antibodies directed against these proteins, and to a vaccine composition comprising these proteins or antibodies directed against these proteins for use in the treatment of infections caused by the *Salmonella spp.* The fimbrial proteins according to the invention or antibodies directed against them can be used for detection of *Salmonella spp.* bacteria.

Background of the invention

The members of genus *Salmonella spp* colonize and infect a wide range of different organisms. Many cause gastroenteritis and enteric fever in humans and domesticated animals while others are not associated with human disease (Saylers et al, 1994). The genus has been divided into two species, *Salmonella bongori* and *Salmonella enterica* where *enterica* can be further subdivided into seven subspecies, designated I, II, IIIa, IIIb, IV, VI, and VII (Reeves et al, 1989). *Salmonella enterica* subspecies I are preferentially associated with warm-blooded animals. Over 99% of all clinical *Salmonella* isolates are strains belonging to this subspecies, including serovars Typhimurium and Enteritidis, which are the major causes of *Salmonella* induced gastroenteritis in humans, and Typhi, the human specific causative organism of typhoid fever, the most severe form of human salmonellosis (Popoff et al, 1992).

Salmonella enterica subspecies I consists of over 1300 different serovars and is preferentially associated with warm-blooded animals (Bäumler, 1997). Over 99% of all clinical *Salmonella* isolates are strains belonging to this subspecies,

including serovars Typhimurium and Enteritidis, which are the major causes of *Salmonella* induced gastroenteritis in humans, and Typhi, the human specific causative organism of typhoid fever, the most severe form of human salmonellosis (Popoff and Le Minor, 1992).

Today gastroenteritis and enteric fever can neither be prevented nor treated with good results. Typhoid fever is a substantial public health problem in developing countries. Each year 33 million people become ill and over 500 000 people die from this infection (American Institute of Medicine, 1986). Typhoid fever can be prevented by vaccination with attenuated bacteria, such as Ty21 and Vi vaccines and whole cell vaccines. Whole cell vaccines show a high incidence of side effects (Ashcroft et al, 1964, Yugoslav Typhoid commission, 1964). The vaccines consisting of attenuated strains of Salmonella typhi suffer from serious drawbacks. They must be administered as three or four spaced doses in order to stimulate protective immune responses (Levine et al, 1989). The treatment of Salmonella typhi with antibiotics is jeopardized since there are strains of Salmonella typhi that are resistant to chloramphenicol, ampicillin, and trimethoprim as well as ciprofloxacin (i.e. multidrug-resistant strains) (Rowe et al, 1997).

Accurate detection of *Salmonella enterica* subspecies I is today not possible. *Salmonella enterica* subspecies I can today only be detected by antibodies directed against surface proteins of *Salmonella enterica* subspecies I. The use of the sequences according to the invention makes it for the first time possible to rapidly and accurately determine the presence of *Salmonella enterica* subspecies I.

For many pathogenic bacteria, there is evidence that the filamentous surface protein structures called pili (fimbriae) are connected to the adhesion of the bacteria to the host cells. Pili proteins are very antigenic and are easily purified. Therefore pili preparations have been used as antigens for vaccination.

Summary of the invention

The invention relates to the objects as defined in the claims. The main object of the present invention is to provide two fimbrial proteins that are specific for *Salmonella enterica* subspecies I bacterial strains, the nucleotide sequences

encoding said proteins, as well as the corresponding amino acid sequences of for therapeutic and diagnostic use. Further are recombinant microorganisms provided, in which the nucleotide sequences according to the invention have been inserted.

An object of the present invention is to provide vaccine compositions for use in the treatment of *Salmonella enterica* infective strains, essentially pure Saf and Tcf fili protein of *Salmonella enterica* subspecies I and *Salmonella enterica* subspecies I serovar Typhi, respectively, as well as antibodies directed to these fili proteins.

A further object of the present invention is to provide the DNA sequences of the genes encoding the Saf and Tcf proteins. These sequences can be used for recombinant production of the proteins and for the preparations of vector vaccines against *Salmonella enterica* subspecies 1 and *Salmonella enterica* subspecies 1 serovar Typhi, respectively, as well as for diagnostic purposes.

Yet another object of the present invention to use purified Saf and Tcf protein from *Salmonella enterica* subspecies 1 bacteria for active or passive immunization of mammals, i.e. the proteins according to the invention can be comprised in a vaccine composition or be used to raise antibodies which can be comprised in a vaccine composition.

Finally, an object of the present invention is to provide a method for preventing or reducing the possibility of *Salmonella* infection of a mammal by administering the vaccines according to the invention. The invention may be more fully understood by reference to the following drawings and detailed description.

Brief description of the drawings

Figure 1.

Schematic representation of phage clones (named N10, D1, B1, F11) covering the entire cs7 insert of *Salmonella enterica* serovar Typhimurium strain SRχ3181, i.e. comprising the *saf* fimbrial operon, i.e. *safA, B, C* and *D* (SEQ ID NO 1).

The clones were selected from partial *Eco* RI and *Bam*HI libraries in the Lambda Dash II vector. The cs7 insert is represented by a bold line. The extent

of respective phage insert is represented by horizontal bars. Name and size of the phage inserts are indicated on the left side of the figure.

Figure 2.

Schematic representation of the pTY52 cosmid comprising the *tcf*-operon (SEQ ID NO 2).

A *tcf* specific PCR fragment of 11105 bp was cloned into the Expand vector I cosmid (Roche). The insert is represented with a thick black line while vector sequences are represented with thin lines. Relevant restriction sites sequences are indicated. The position of the *tcf*-operon, i.e. *tcfA, B, C* and *D* (SEQ ID NO 2), is represented by a shaded arrow.

Figure 3.

The phylogenetic distribution of the identified genes on the cs7 insert was investigated using the well defined SARC collection, see Example 1.

Figure 4.

A 2 kb large internal *Eco*R I fragment was used as a probe in a Southern blot of the SARC collection, see Example 2.

Sequence listing

SEQ ID NO 1—DNA sequence of the genes encoding the precursor of the saf fimbrie unit of *Salmonella enterica* subspecies I.

SEQ ID NO 2—DNA sequence of the genes which encode the precursor of the tcf fimbrie unit of *Salmonella enterica* subspecies I serovar Typhi.

Deposit information

The phages carrying the inserted SEQ ID NO 1, i.e. phages clones B1, D1, F11 and N10 (see Figure 1) have been given the ECACC Accession numbers 99051922, 99051923, 99051924, and 99051925, respectively.

The cosmide carrying the inserted SEQ ID NO 2, i.e. cosmide pTY52 (see Figure 2) has been given the ECACC Accession number 99051926.

The depositions were made May 19, 1999.

Detailed description of the invention

The present invention is based on the finding that two fimbrial operons, the *saf* operon and the *tcf* operon, are specific for *Salmonella enterica* subspecies 1 bacteria. Due to their specificity they can be used to provide vaccines against *Salmonella enterica* subspecies I as well as detection methods for *Salmonella enterica* subspecies I. The *saf* operon is specific for all *Salmonella enterica*

subspecies 1 bactéria and the *tcf* operon is specific for the serovar typhi of *Salmonella enterica* subspecies 1, see Examples 1 & 2.

The main object of the invention relates to two fimbrial operons, the *saf* operon and the *tcf* operon, that are specific for *Salmonella enterica subspecies* 1 bacteria for terapeutic use.

Another object of the present invention is to provide vaccines against *Salmonella enterica* subspecies 1 induced gastroentritis, entric fever and typhoid fever.

A further object of the present invention is to provide methods to detect *Salmonella enterica* subspecies 1. The nucleotide sequences according to the invention are useful for constructing vectors for use as vaccines for insertion into attenuated bacteria in constructing a recombinant vaccine, for insertion into a viral vector in constructing a recombinant viral vaccine, or for direct inoculation as a nucleic acid vaccine. The pili proteins according to the invention, or antigenic fragments thereof, can be used for active immunization and antibodies directed against them can be used for passive immunization. All these applications of the sequences according to the invention are obtained by applying standard techniques known to the man ordinary skilled in the art.

Vaccines against *Salmonella enterica* subspecies I.
The genes encoding the saf and tcf fimbrial structures, or fragments thereof, may be incorporated into a bacterial or viral vaccine comprising recombinant bacteria, virus or fungi which are engineered to produce one or more immunogenic epitopes of the saf or tcf fimbrial structures. In addition, the genes encoding the saf and tcf fimbrial structures, or part thereof, operatively linked to regulatory elements, can be introduced directly as a nucleic acid vaccine, to elicit a protective immune response.

The proteins or antigenic fragment thereof, deduced from the nucleic acid sequences of the present invention are useful alone or in conventional vaccine mixtures in the vaccine compositions according to the invention. The proteins could be produced by chemical synthesis or recombinant expression according to conventional methods.

The proteins and peptides according to the invention can be obtained by using a host organism transformed or transfected with an expression vector obtained by insertion of a gene according to the invention, or part thereof, into a vector in a conventional manner. The vector which is used to construct the expression vector is not particularly limited, but specific examples include plasmids such as pET (Stratagen) and the like; and phages such as M13 (NEB), phage display libraries and the like. As expression regulatory sequence can among others T7 promotors and lac promotors be used.

An appropriate host to be transformed or transfected with the expression vector can be chosen among for example *E.-coli, Salmonella* or *Bacillus subtilus*. The transformed or transfected host is cultured and proliferated under suitable conditions.

After culturing, the peptides of the present invention may be purified by, for example, chromatography, precipitation, and/or density gradient centrifugation. The thus obtained peptides can be used as a vaccine or for the production of antibodies directed against said peptides, which can be used for passive immunization.

The purified preparation containing one or several proteins according to the invention, or parts thereof, is then formulated as a pharmaceutical composition, as for example a vaccine, or in a mixture with adjuvants. If desired the proteins are fragmented by standard chemical or enzymatic techniques to produce antigenic segments.

In formulating the vaccine compositions with the peptide or protein, alone or in various combinations, the immunogen is adjusted to an appropriate concentration and formulated with any suitable vaccine adjuvant. The immunogen may also be incorporated into liposomes, or conjugated to polysaccharides and/or other polymers for use in a vaccine formulation.

The different vaccines according to the present invention are administered to mammals in many different ways. These include intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, oral, and intranasal routes of administration. The vaccine doses will differ depending on circumstances such

as body weight, interferences with other administered medicaments etc. The upper limit is not critical unless the dose shows toxicity.

The peptides and proteins of the present invention are also useful to produce monoclonal or polyclonal antibodies for use in passive immunotherapy against *Salmonella enterica* subspecies 1. Human immunoglobulin is preferred. Antisera is obtained from individuals immunized with proteins or peptides according to the invention. The immunoglobulin fraction is then enriched, for example by immunoaffininty or affininty chromatography. Antibodies raised in a suitable mammal or in the patient to be treated, can subsequently be administered locally or topically, e.g. orally to the patient.

Detection of *Salmonella enterica* subspecies I in general.
The sequences according to the invention, or part thereof, or fragments hybridizing therewith, as well as the proteins according to the invention, or part thereof, and antibodies directed to said proteins, or antigenic fragments thereof, can be used in molecular diagnostic assays for the detection of *Salmonell enterica* subspecies I.

Nucleic acids having the nucleotide sequence according to the invention, or any nucleotide sequence hybridizing therewith can be used as a probe in nucleic acid hybridization assays for the detection of *Salmonella spp* in various tissues and body fluids of patients. The hybridization assay may be of any type including; Southern blots, Northern blots, colony blots.

PCR technology is the most preferred technology for detection according to the invention of *Salmonella enterica* subspecies 1. Primers of at least one selected from the 5' end and one from the 3' end can be used in PCR and other known tests to rapidly identify the presence of *Salmonella enterica* subspecies 1. This is according to conventional techniques.

The isolated and purified proteins and peptides of the invention can be used as diagnostics to measure an increase in serum titer of *Salmonella enterica* subspecies I-specific antibody since they bind strongly to these antibodies. A serum test sample can be screened for *Salmonella enterica* subspecies I by methods such as for example ELISA.

The invention further comprises the use of antibodies directed against the saf and tcf fimbrie structures for quantitative or qualitative determinations of the pili proteins of the invention, or fractions thereof, in cells, tissues or body fluids.

Detection of *Salmonella enterica* subspecies I by using nucleic acid hybridization technology

Nucleic acid hybridization technology can also be to detect *Salmonella enterica* subspecies 1 according to the invention. The nucleic acid probes chosen from parts of the sequences according to the invention can be either DNA or RNA. DNA sequences complementary to the sequences according to the invention can also be used. The binding of the probe to the target sequence, i.e. the hybridization, must not be perfect. Variations and mutations of the sequences according to the invention can be used as long as they hybridize good enough to detect *Salmonella enterica* subspecies I. The preferred length of the nucleic acid probes is about 10 to 400 nucleotides, most preferred not longer than 100 nucleotides.

The nucleotide probe is preferably chosen from the parts of the sequences that have the least variation. In the most preferred embodiments when screening for SEQ ID NO 1 (the *saf* operon, specific for *Salmonella enterica* subspecies 1) a nucleotide probe or PCR primer selected from nucleotides 37 368-37 868 should be avoided since this region is hypervariable.

The nucleic acid probes according to the invention are prepared by any conventional method such as organic synthesis, recombinant techniques, or isolation from genomic DNA.

The nucleic acid probes of the invention are labeled in a conventional manner to signal hybridization to target nucleic acid from *Salmonella enterica* subspecies I. The labeling may comprise a radiolabel, an enzyme, a bacterial label, a fluorescent label, an antibody, an antigen, a latex particle, an electron dense compound, or a light scattering particle.

The probes may be provided in a lyophilized form, to be reconstituted in a buffer appropriate for hybridization, or the probes may already be present in

such a buffer. The buffer may contain a suitable hybridization enhancer, detergent, carrier DNA, and a compound to increase the specificity.

Any conventional hybridization assay technique, such as dot blot hybridization, Southern blotting, sandwich hybridization, displacement hybridization and the like, can be used.

The target analyte polynucleotide of a microorganism may be in various media, most often in a biological, or physiological specimen. In most cases it is preferred to subject the specimen containing the target polynucleotide to any conventional extraction, purification, and/or isolation before conducting the analysis.

The sample containing the target analyte nucleotide sequence must often be treated to convert the DNA to a single-stranded form, which may be accomplished by a variety of conventional techniques, such as thermal or chemical techniques.

The following examples describe the isolation and specificity of the sequences according to the invention.

EXAMPLE 1

Identification and characterization of the *saf* operon.

The present inventors found, upon investigation of a 7 kb chromosomal region on centisome 7 originally isolated from the *S. typhimurium* strain SR-11 $_k$3181, a region that exhibits many of the traits that define a pathogenicity island. It has a lower G+C composition than the average composition of the *Salmonella* genome and includes many sequences related to different mobile genetic elements. The region is not present in *E.coli* K12, and the *Salmonella* specific DNA is inserted between the tRNA gene *aspV* and the stop codon of *yafV,* a hypothetical protein upstream of the *yafH* gene at 5 min in the *E.coli* chromosome. This *Salmonella* specific insert encodes proteins creating adhesive structures and other virulence factors. Sequencing revealed genes encoding a new fimbrial operon that they designated *Salmonella* Atypical Fimbriae (*saf*), due to its relatedness to a subgroup of adhesive structures forming thin atypical fimbriae or non-fimbrial adhesins.

The *saf* operon consists of four contiguous genes, *safA*, *safB*, *safC* and *safD* that encode fimbrial subunit, periplasmic chaperone, outer membrane usher protein and alternative fimbrial subunit, respectively. The genes *safA, B, C* and *safD* encode putative proteins of 166, 244, 836 and 156 amino acids, respectively. Analyzes of clinical *Salmonella* isolates showed that DNA of 195 out of 198 clinical isolates belonging to *S. enterica* subspecies I hybridized with *safB* and *safC,* i.e. these sequences are common to more than 99% of the known *Salmonella enterica* subspecies 1 bacteria. The inventors showed that 58% of these clinical isolates carry the *safA*, see Table 1.

Table 1. The prevalence of the *saf* genes in clinical *Salmonella* isolates.

Serovar	*safA*	*safB*	*safC*	# isolates
S. adelaide	-	+	+	1
S. agona	+	+	+	6
S. anatum	-	+	+	3
S. bareilly	+	+	+	3
S. blockley	+	+	+	3
S. bovismorbificans	-	+	+	5
S. braenderup	-	+	+	4
S. brandenburg	+	+	+	1
S. bredeney	+/-	+	+	15
S. chester	+	+	+	1
S. colindale	-	+	+	1
S. derby	-	+	+	1
S. dublin	-	+	+	1
S. eastbourne	+	+	+	2
S. emek	+	+	+	1
S. enteritidis	-	+	+	8
S. give	-	+	+	1
S. goettingen	+	+	+	1
S. haardt	-	+	+	1
S. hadar	+	+	+	16
S. heidelberg	-	+	+	1
S. hvittingfoss	+	+	+	5
S. infantis	-/+	+	+	6
S. java	-	+	+	1
S. javiana	-	+	+	1
S. kottbus	-	+	+	1
S. livingstone	-	+	+	1
S. london	+	+	+	1
S. maastricht	+	+	+	2
S. mbandaka	-	-	-	3
S. montevideo	+	+	+	1
S. muenster	-	+	+	1
S. newport	+	+	+	2
S. ohio	+	+	+	1
S. oranienburg	+	+	+	2
S. panama	+	+	+	3
S. potsdam	+	+	+	1
S. rissen	-	-	-	1
S. saarbrucken	-	+	+	1
S. saint paul	+	+	+	3
S. schwartzengrund	-	+	+	1
S. singapore	+	+	+	1
S. stanley	+	+	+	5
S. subsp I 4.5,12:-:-	+	+	+	2
S. subsp I 4.5,12:b:-	-	+	+	1
S. subsp I 4.5,12:i:-	+	+	+	1
S. subsp I spont	-	+	+	1
S. tennessee	+	+	+	2
S. thompson	-	+	+	1
S. typhi	-	+	+	1
S. typhimurium	+	+	+	27
S. virchow	+	+	+	7
S. weltervreden	-	+	+	1
S. worthington	-	-	-	2
S. subsp III	-	-	-	1

The phylogenetic distribution of the identified genes on the cs7 insert was investigated using the well defined SARC collection, which showed that the presence of the *safA, safB, safC* and *safD* genes is restricted to *S. enterica* subspecies I (Fig. 3). This region is hence the first subspecies I specific genetic region to be identified with a broad distribution within the subspecies. Since the serovars of subspecies I constitute over 99% of human salmonellosis and are preferentially associated with warm blooded animals, it implicates a role for the *saf* adhesive organelle in the colonization of these organisms.

EXAMPLE 2

Identification and characterization of the *tcf* operon.

The present inventors found that *Salmonella enterica* subspecies I serovar Typhi contains DNA encoding an additional fimbrial operon, the tcf operon, in the *sinR-pagN* intergenic region. Southern blot analysis revealed a markedly different restriction pattern in *S. enterica* serovar Typhi than the other subspecies I isolates, suggesting that the *saf-sin* region in serovar Typhi might carry additional DNA relative to serovar Typhimurium strains. A PCR reaction (using a kit from Roche) was therefore performed using a *sinR* (5'-GTA AAT CGC TTA GTC GCC-3') specific forward primer and a *pagN* (5'-TCA ACT CAA CCT TCA GCC-3') specific reverse primer.

This primer pair produced, as expected, a product of 2 kb in serovar Typhimurium from the SARC collection, while from serovar Typhi the product was 10 kb. Thus, the neighboring *sinR* and *pagN* genes in serovar Typhimurium strains are separated by approximately 8 kb in serovar Typhi.

The Typhi specific PCR product was purified, digested partially with *EcoRI* and sub-cloned into pUC18 forming a set of overlapping clones. Sequencing of the clones revealed a putative fimbrial operon designated *tcf* for Typhi Colonizing Factor. Four ORFs, *tcfA,B,C,D*, have been identified with putative proteins having significant homology to CooB (38% identical over 192 aa), CooA (37% identical over 170 aa), CooC (34% identical over 872 aa) and CooD (31% identical over 272 aa), respectively. The Coo proteins are involved in the biosynthesis of the CS1 colonizing factor antigens of enterotoxigenic *E.coli* (Fig. 4) (Froehlich et al., 1994). The peptide of the *tcfB* ORF is also homologous to the CblA major fimbrial subunit protein (45% identical over 154 aa) of the cable

type II pili of the cystic fibrosis-associated *Burkholderia cepacia*(Sajjan et al., 1995). Down-stream of the *tcf*-operon two ORFs were identified with the same transcriptional orientation as the *tcf* genes. The first was designated *tinR* for Typhi insert regulator because it is homologous (33% identical over 144 aa) to AzlB of *Bacillus subtilis*, a member of the Lrp/AsnC family of transcriptional regulators (Belitsky et al., 1997). *tinR* is followed by an ORF (*tioA* for Typhi insert orf) encoding a putative protein of 205 amino acids with no significant homologies to anything in the DDBJ/EMBL/GenBank databases. The above sequence from *Salmonella enterica* serovar Typhi strain RKS 3333 and the *tcf* region of the incomplete genome sequence from serovar Typhi strain CT18 (http:// www.sanger.ac.uk) are 99% identical over the total length of the investigated region in concordance with the clonal nature of the serovar .

A 2 kb large internal *Eco*R I fragment was used as a probe in a Southern blot of the SARC collection. This blot shows that *Salmonella enterica* subspecies I serovar Typhi (SARC2) is the only strain in the collection possessing DNA hybridizing to this fragment (Fig. 4).